SecureFinAI Contest Task 1: FinRL-Transformer for Cryptocurrency Trading

Laxmi Tiwari

Logictronix Technologies

Objectives

Develop an offline reinforcement learning system for high-frequency Bitcoin trading using Decision Transformers (DT).

- Leverage BTC order book data and sentiment analysis for robust trading strategies.
- Model sequences of market states and returns to predict optimal actions: Buy, Hold, Sell.
- Train using offline RL trajectories for safe and data-efficient learning.

Introduction

Decision Transformers (DT) combine sequence modeling with reinforcement learning by framing RL as a supervised learning problem. They predict actions conditioned on past states and **return-to-go**, enabling goal-directed behavior in episodic environments like crypto trading.

This approach bypasses unstable online exploration and leverages historical trajectories to learn effective policies efficiently.

Highlights

- Offline RL using historical BTC data and sentiment features.
- Return-to-go conditioning for goal-directed trading.
- Causal attention ensures proper sequential modeling.
- Interpretable evaluation using trajectory visualizations and portfolio metrics.

State Representation

The state representation is designed for crypto-specific dynamics, integrating LOB, on-chain, and sentiment features beyond traditional stock-based alpha factors.

Methodology

- **Data:** 1-second BTC order book + sentiment and risk scores.
- **Preprocessing:** Generate New Technical Factors, normalize sequences.
- **RL Trajectories:** Train PPO/DQN agents and convert replay buffers to DT format.
- **Decision Transformer:** Context length = 20, hidden size = 64, 3 transformer layers, 1 attention head.
- **Training:** Supervised learning on trajectory dataset, predicting next action conditioned on state and return-to-go.
- **Evaluation:** Portfolio metrics including cumulative return, Sharpe ratio, max drawdown, and win rate.

Architecture Overview

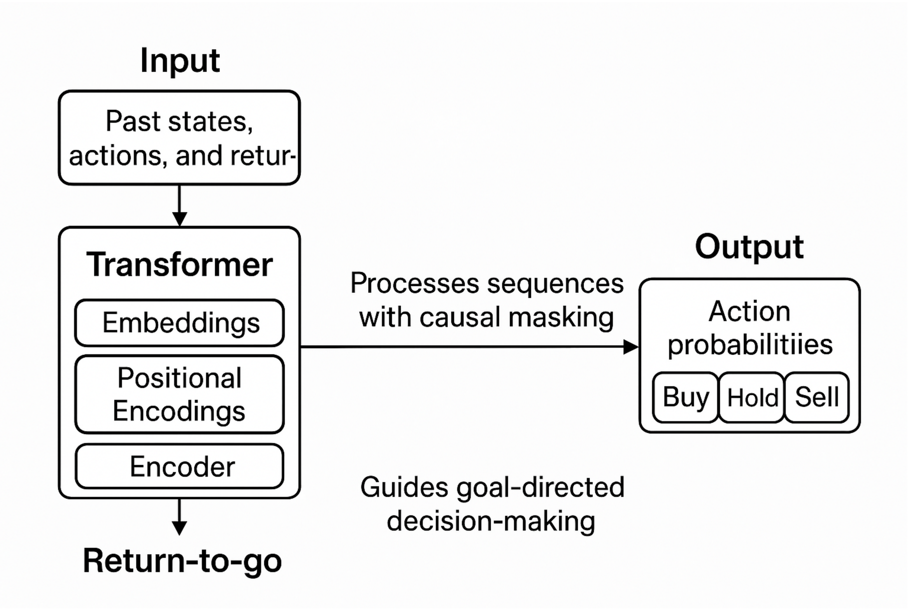

Figure 1:Decision Transformer-based Trading Architecture

- Input: Past states, actions, and returns.
- Transformer: Processes sequences with causal masking.
- Output: Action probabilities (Buy, Hold, Sell).
- Return-to-go: Guides goal-directed decision-making.

Key Takeaways

- Decision Transformers successfully learn trading policies from offline BTC order book and sentiment data.
- Dilated CNN models aggregate sequences of new technical indicators specifically designed for LOB data in crypto, replacing traditional weak factors like Alpha 101.
- Return-to-go conditioning enables goal-directed behavior, allowing the agent to target specific portfolio outcomes.
- Offline training reduces the risk of live market exploration while still capturing temporal dependencies in high-frequency data.
- The model demonstrates superior performance compared to baseline PPO and DQN agents in cumulative returns, Sharpe ratio, and drawdown.
- Trajectory visualizations and portfolio metrics provide interpretable insights, making the agent's decisions more transparent and trustworthy.

Findings

Using **Decision Transformers** enhances **offline learning**, making a significant impact on advancing **trading research** and data-driven strategy development.

Learning Curve of model

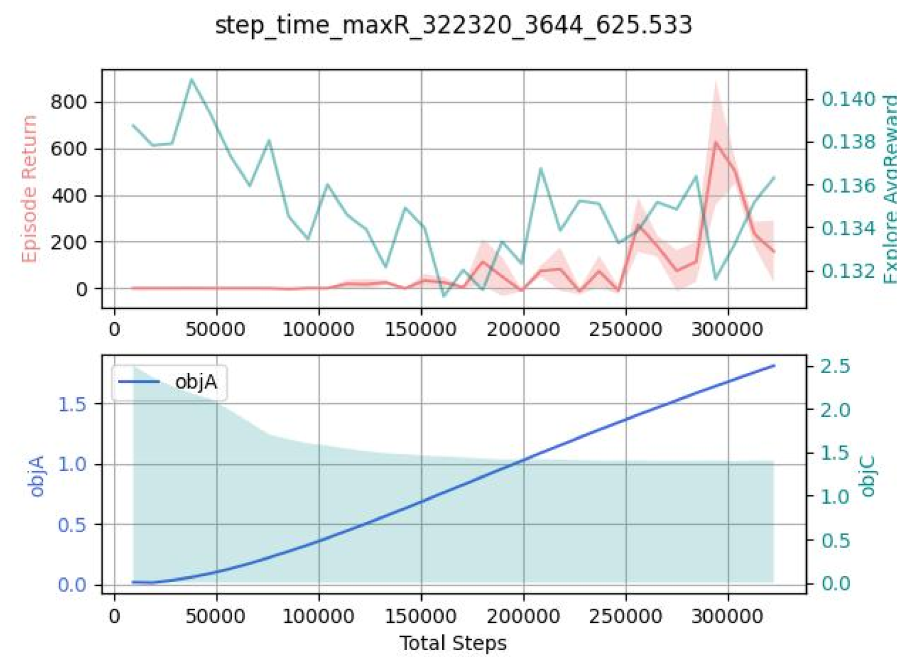

Figure 2:Learning Curve of the Model

Conclusion

Decision Transformers provide a **robust, goal-directed framework** for offline crypto trading:

- Learns from historical trajectories without risky online exploration.
- Captures temporal dependencies in BTC market data.
- Produces interpretable policies with clear action trajectories.
- Demonstrates superior portfolio metrics compared to standard PPO/DQN agents.

Acknowledgements

We sincerely acknowledge Logictronix Technologies. Their guidance, resources, and insights were instrumental in this project.

Contact Information

- Email: laxmi@logictronix.com

logictronix